PATENT DOCKET NO. PR1237
EXPRESS MAIL NO: EM 168884885 US
MAILED: March 25, 1998

Neurotrimin Homolog

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides which are characterized by the presence of characteristics particular to a novel subfamily of immunoglobulin-like (Ig) glycosyl-phosphatidylinositol (GPI)-anchored cell adhesion molecules (CAMs) proteins. Specifically, this subfamily is known as IgLON and includes neurotrimin, opiate binding cell adhesion molecule (OBCAM), limbic system-associated membrane protein (LAMP), CEPU-1, GP-55 and AvGp50.

BACKGROUND OF THE INVENTION

Neuronal development in higher vertebrates is characterized by processes that must successfully navigate distinct cellular environment en route to their synaptic targets. The result is a functionally precise formation of neural circuits. The precision is believed to result form mechanisms that regulate growth cone pathfinding and target recognition, followed by latter refinement and remodeling of such projections by events that require neuronal activity, Goodman and Shatz, *Cell/Neuron [Suppl.]* 72(10): 77-98 (1993). It is further evident that different neurons extend nerve fibers that are biochemically distinct and rely on specific guidance cues provided by cell-cell, cell-matrix, and chemotrophic interactions to reach their appropriate synaptic targets, Goodman *et al., supra.*

One particular means by which diversity of the neuronal cell surface may be generated is through differential expression of cell surface proteins referred to as cell adhesion molecules (CAMs). Neuronally expressed CAMs have been implicated in diverse developmental processes, including migration of neurons along radial glial cells, providing permissive or repulsive substrates for neurite extension, and in promoting the selective fasciculation of axons in projectional pathways. Jessel, *Neuron* 1: 3-13 (1988); Edelman and Crossin, *Annu. Rev. Biochem.* 60: 155-190 (1991). Interactions between CAMs present on the growth cone membrane and molecules on opposing cell membranes or in the extracellular matrix are thought to provide the specific guidance cues that direct nerve fiber outgrowth along appropriate projectional pathways. Such interactions are likely to result in the activation of various second messenger systems within the growth cone that regulate neurite outgrowth. Doherty and Walsh, *Curr. Opin Neurobiol.* 2: 595-601 (1992).

In higher vertebrates, most neural CAMs have been found to be members of three major structural families of proteins: the integrins, the cadherins, and the immunoglobulin gene superfamily (IgSF). Jessel, *supra.*; Takeichi,

Annu. Rev. Biochem. 59: 237-252 (1990); Reichardt and Tomaselli, *Annu. Rev. Neurosci.* 14: 531-570 (1991). Cell adhesion molecules of the IgSF (or Ig-CAMs), in particular, constitute a large family of proteins frequently implicated in neural cell interactions and nerve fiber outgrowth during development, Salzer and Colman, *Dev. Neurosci.* 11: 377-390 (1989); Brümmendorf and Rathjen, *J. Neurochem.* 61: 1207-1219 (1993). However, the majority of mammalian Ig-CAMs appear to be too widely expressed to specify navigational pathways or synaptic targets suggesting that other CAMs, yet to be identified, have role in these more selective interactions of neurons.

Many of the known neural Ig-CAMs have been found to be attached to the plasma membrane via a glycosylphosphatidylinositol (GPI) anchor. Additionally, many studies have implicated GPI-anchored proteins in providing specific guidance cues during the outgrowth on neurons in specific pathways. In studies of the grasshopper nervous system, treatment of embryos with phosphatidylinositol-specific phopholipase C (PIPLC), which selectively removes GPI-anchored proteins from the surfaces of cells, resulted in misdirection and faulty navigation among subsets of pioneering growth cones, as well as inhibited migratory patterns of a subset of early neurons, Chang *et al., Devel.* 114: 507-519 (1992). The projection of retinal fibers to the optic tectum appears to depend, in part, on a 33 kDa GPI-anchored protein, however, the precise nature of this protein is unknown. Stahl *et al., Neuron* 5: 735-743 (1990).

The expression of various GPI-anchored proteins has been characterized amongst the different populations of primary rat neurons amongst dorsal root ganglion, sympathetic neurons of the cervical ganglion, sympathetic neurons of the superior cervical ganglion, and cerebellar granule neurons. Rosen *et al., J. Cell Biol.* 117: 617-627 (1992). In contrast to the similar pattern of total membrane protein expression by these different types of neurons, striking differences were observed in the expression of GPI-anchored proteins between these neurons. Recently, a 65 kDa protein band known as neurotrimin was discovered and found to be differentially expressed by primary neurons (Rosen *et al., supra*), and restricted to the nervous system and found to be the most abundant and earliest expressed of the GPI-anchored species in the CNS. Struyk *et al., J. Neuroscience* 15(3): 2141-2156 (1995). The discovery of neurotrimin has further lead to the identification of a family of IgSF members, each containing three Ig-like domains that share significant amino acid identity, now termed IgLON. Struyk *et al., supra*; Pimenta *et al., Gene* 170(2): 189-95 (1996).

Additional members of the IgLON subfamily include opiate binding cell adhesion molecule (OBCAM), Schofield *et al., EMBO J.* 8: 489-495 (1989); limbic associated membrane protein (LAMP), Pimenta *et al., supra*; CEPU-1; GP55, Wilson *et al., J. Cell Sci.* 109: 3129-3138 (1996); *Eur. J. Neurosci.* 9(2): 334-41 (1997); and AvGp50, Hancox *et al., Brain Res. Mol. Brain Res.* 44(2): 273-85 (1997).

While the expression of neurotrimin appears to be widespread, it does appear to correlated with the development of several neural circuits. For example, between E18 and P10, neurotimin mRNA expression within the forebrain is maintained at high levels in neurons of the developing thalamus, cortical subplate, and cortex, particularly laminae V and VI (with less intense expression in II, II, and IV, and minimal expression in lamina I). Cortical subplate neurons may provide an early, temporary scaffold for the ingrowing thalamic afferents en route to their final synaptic targets in the cortex. Allendoerfer and Shatz, *Annu. Rev. Neurosci.* 17: 185-218 (1994).

Conversely, subplate neurons have been suggested to be required for cortical neurons from layer V to select VI to grow into the thalamus, and neurons from layer V to select their targets in the colliculus, pons, and spinal cord (McConnell *et al., J. Neurosci.* 14: 1892-1907 (1994). The high level expression of neurotrimin in many of these projections suggests that it could be involved in their development.

In the hindbrain, high levels of neurotrimin message expression were observed within the pontine nucleus and by the internal granule cells and Purkinje cells of the cerebellum. The pontine nucleus received afferent input from a variety of sources including corticopontine fibers of layer V, and is a major source of afferent input, via mossy fibers, to the granule cells which, in turn, are a major source of afferent input via parallel fibers to Purkinje cells. [Palay and Chan-Palay, *The cerebellar cortex: cytology and organization.* New York: Springer (1974]. High level expression of neurotrimin these neurons again suggests potential involvement in the establishment of these circuits.

Neurotrimin also exhibits a graded expression pattern in the early postnatal striatum. Increased neurotrimin expression is found overlying the dorsolateral striatum of the rat, while lesser hybridization intensity is seen overlying the ventromedial striatum. Struyk *et al., supra.* This region of higher neurotrimin hybridization intensity does not correspond to a cytoarchitecturally differentiable region, rather it corresponds to the primary area of afferent input from layer VI of the contralateral sensorimotor cortex (Gerfen, *Nature* 311: 461-464 (1984); Donoghue and Herkenham, *Brain Res.* 365: 397-403 (1986)). The ventromedial striatum, by contrast, receives the majority of its afferent input from the perirhinal and association cortex. It is noteworthy that a complementary graded pattern of LAMP expression, has been observed within the striatium, with highest expression in ventromedial regions, and lowest expression dorsolaterally. Levitt, *Science* 223: 299-301 (1985); Chesselet *et al., Neuroscience* 40: 725-733 (1991).

SUMMARY OF THE INVENTION

Applicants have identified a cDNA clone (DNA43316) that encodes a novel polypeptide, designated in the present application as "PRO337."

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO337 polypeptide comprising the sequence of amino acids 1 to 1795 of Fig. 1 (SEQ ID NO: 1), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95 (including 96, 97, 98 and 99%) sequence identity with a polypeptide having amino acid residues 1 to 1795 of Fig. 1 (SEQ ID NO: 1). Preferably, the highest degree of sequence identity occurs within the immunoglobulin and major histocompatibility domains (amino acids 113 to 130 of Fig. 2, SEQ ID NO: 2). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a neurotrimin polypeptide having amino acid residues 1 to 344 of Fig. 2 (SEQ ID NO: 2), or is complementary to such encoding nucleic acid

sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA43316-1237, deposited with the ATCC under accession number ATCC 209487, alternatively the coding sequence of clone DNA43316-1237, deposited under accession number ATCC 209487.

In yet another embodiment, the invention provides a vector comprising DNA encoding PRO337 polypeptide. A host cell comprising such a vector is also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing PRO337 polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of PRO337 and recovering the same from the cell culture.

In yet another embodiment, the invention provides isolated PRO337 polypeptide. In particular, the invention provides isolated native sequence PRO337 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 344 of Figure 2 (SEQ ID NO: 2). Native PRO337 polypeptides with or without the native signal sequence (amino acids 1 to about 28 in Figure 2 (SEQ ID NO: 2), and with or without the initiating methionine are specifically included. Alternatively, the invention provides a PRO337 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209487.

In yet another embodiment, the invention provides chimeric molecules comprising a PRO337 polypeptide fused to a heterologous polypeptide or amino acid sequence. An example of such a chimeric molecule comprises a PRO337 polypeptide fused to an epitope tag sequence or an Fc region of an immunoglobulin.

In yet another embodiment, the invention provides an expressed sequence tag (EST) comprising the nucleotide sequences identified in Fig. 3 as DNA28761 (SEQ ID NO: 3), HSC0PH101 (SEQ ID NO: 4), R18841 (SEQ ID NO: 5) and 1333422 (SEQ ID NO: 6). Additionally, the invention provides an expressed sequence tag (EST) comprising the nucleotide sequence identified in Fig. 9 as DNA42301.

In yet another embodiment, the invention provides an antibody which specifically binds to PRO337 polypeptide. Optionally, the antibody is a monoclonal antibody.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the nucleotide sequence of a native sequence DNA43316 cDNA.

Figure 2 shows the derived amino acid sequence of a native sequence PRO337 polypeptide.

Figure 3 shows the alignment comparison between DNA28761 (fromDNA) and further sequences from the (LIFESEQ™ database, Incyte Pharmaceuticals, Palo Alto, CA and GenBank), which were used to extend the from DNA to obtain a consensus sequence shown in the bottom line each cluster of the Figure.

Figure 4 shows the oligonucleotide sequences OLI620 (28761.p), OLI621 (28761.f) and OLI622 (28761.r) which were used in the isolation of DNA43316.

Figure 5 describes the Blast score, match and percent homology alignment between the coded protein of DNA43316 and certain other members of the IgLON immunoglobulin superfamily.

Figure 6 shows the double stranded nucleotide sequence corresponding to DNA43316, the amino acid sequence encoded thereby (PRO337) and the cloning oligos depicted in Figure 4.

Figure 7 identifies the major domains and homologies which are present in PRO337. Fig. 7A is a sequence of PRO337 with corresponding residue identification. Fig. 7B identified the signal sequence and immunoglobulin and major histocompatibility domains.

Figure 8A is a alignment comparison between human neurotrimin (DNA43316) and other members of the IgLOM, LAMP, OBCAM, CEPU-1 and AvGp50 families. Figure 8B is a tree comparison of the sequence identities.

Figure 9 is a comparison of the EST sequence DNA42301 with DNA43316.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "PRO337" " PRO337 polypeptide" and "human neurotrimin" when used herein encompass native sequence PRO337 and PRO337 variants (which are further defined herein). PRO337 may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

A "native sequence PRO337" comprises a polypeptide having the same amino acid sequence as a PRO337 derived from nature. Such native sequence PRO337 can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO337" specifically encompasses naturally-occurring truncated or secreted forms of PRO337 (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of PRO337. In one embodiment of the invention, the native sequence PRO337 is a mature or full-length native sequence human neurotrimin comprising amino acids 1 to 344 of Fig. 2 (SEQ ID NO: 2), with or without the N-terminal signal sequence (residues 1 to about 28), and with or without the initiating methionine at position 1.

"PRO337 variant" means an active PRO337 as defined below having at least about 80% amino acid sequence identity to (a) a DNA molecule encoding a PRO337 polypeptide, with or without its native signal sequence, or (b) the complement of the DNA molecule of (a). In a particular embodiment, the PRO337 variant has at least about 80% amino acid sequence homology with the PRO337 having the deduced amino acid sequence shown in Fig. 2 (SEQ ID NO: 2) for a full-length native sequence PRO337. Such PRO337 variants include, for instance, PRO337 polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the sequence of Fig. 2 (SEQ ID NO: 2). Preferably, the nucleic acid or amino acid sequence identity is at least about 85%, more preferably at least about 90%, and even more preferably at least about 95%, including 96%, 97%, 98% and 99%.

"Percent (%) amino acid sequence identity" with respect to the PRO337 sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the PRO337 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST,

BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Preferably, the alignment is determined by BLAST-2 where the identity parameters are set to the default settings.

"Percent (%) nucleic acid sequence identity" with respect to the DNA43316 sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the DNA43316 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Preferably, the alignment is determined by BLAST-2 where the identity parameters are set to the default settings.

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO337 natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" DNA43316 nucleic acid molecule is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the DNA43316 nucleic acid. An isolated DNA43316 nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated DNA43316 nucleic acid molecules therefore are distinguished from the DNA43316 nucleic acid molecule as it exists in natural cells. However, an isolated DNA43316 nucleic acid molecule includes DNA43316 nucleic acid molecules contained in cells that ordinarily express DNA43316 where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it

is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers single anti-PRO337 monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies) and anti-PRO337 antibody compositions with polyepitopic specificity. The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Active" or "activity" for the purposes herein refers to form(s) of PRO337 which retain the biologic and/or immunologic activities of native or naturally-occurring PRO337. A preferred activity is the ability to bind to and affect, e.g., block or otherwise modulate, an activity of neurotrimin. The activity preferably involves the regulation neuron growth and differentiation.

II. Compositions and Methods of the Invention

A. Full-length human neurotrimin

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO337. In particular, Applicants have identified and isolated cDNA encoding a PRO337 polypeptide, as disclosed in further detail in the Examples below. Using BLAST, BLAST-2 and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO337 (shown in Figure 2 and SEQ ID NO: 2) has 97% amino acid sequence identity with rat neurotrimin, 85% sequence identity with chicken CEPU, 73% sequence identity with chicken G55, 59% homology with human LAMP and 84% homology with human OPCAM. Accordingly, it is presently believed that PRO337 disclosed in the present application is a newly identified member of the IgLON sub family of the immunoglobulin superfamily and may possess neurite growth and differentiation potentiating properties.

B. PRO337 Variants

In addition to the full-length native sequence PRO337 described herein, it is contemplated that PRO337 variants can be prepared. PRO337 variants can be prepared by introducing appropriate nucleotide changes into the PRO337 DNA, or by synthesis of the desired PRO337 polypeptides. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO337, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO337 or in various domains of the PRO337 described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative

mutations set forth, for instance, in U.S. Patent No. 5,364,934. . Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO337 that results in a change in the amino acid sequence of the PRO337 as compared with the native sequence PRO337. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO337. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO337 with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity in the *in vitro* assay described in the Examples below.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO337 variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

C. Modifications of PRO337

Covalent modifications of PRO337 are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of the PRO337 with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO337. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO337 to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO337 antibodies, and vice-versa. Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis-(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding

glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α-amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO337 polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO337, and/or adding one or more glycosylation sites that are not present in the native sequence PRO337, and/or alteration of the ratio and/or composition of the sugar residues attached to the glycosylation site(s).

Addition of glycosylation sites to the PRO337 polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO337 (for O-linked glycosylation sites). The PRO337 amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO337 polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO337 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO337 polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO337 comprises linking the PRO337 polypeptide to one of a variety of nonproteinaceous polymers, *e.g.*, polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO337 of the present invention may also be modified in a way to form a chimeric molecule comprising PRO337 fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of the PRO337 with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO337. The presence of such epitope-tagged forms of the PRO337 can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO337 to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO337 with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region

of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α-tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

D. Preparation of PRO337

The description below relates primarily to production of PRO337 by culturing cells transformed or transfected with a vector containing PRO337 nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO337. For instance, the PRO337 sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO337 may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO337.

1. Isolation of DNA Encoding PRO337

DNA encoding PRO337 may be obtained from a cDNA library prepared from tissue believed to possess the PRO337 mRNA and to express it at a detectable level. Accordingly, human PRO337 DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO337-encoding gene may also be obtained from a genomic library or by oligonucleotide synthesis.

Libraries can be screened with probes (such as antibodies to the PRO337 or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO337 is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ^{32}P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are

provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined through sequence alignment using computer software programs such as BLAST, BLAST-2, ALIGN, DNAstar, and INHERIT which employ various algorithms to measure homology.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO337 production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of transfection are known to the ordinarily skilled artisan, for example, $CaPO_4$ and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, *e.g.*, polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635).

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO337-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism.

Suitable host cells for the expression of glycosylated PRO337 are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as Drosophila S2 and Spodoptera Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (*e.g.*, cDNA or genomic DNA) encoding PRO337 may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The PRO337 may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO337 DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α-factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2: plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning

vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO337 nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp*1 gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp*1 gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO337 nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β-lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO337.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO337 transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, *e.g.*, the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the

host cell systems.

Transcription of a DNA encoding the PRO337 by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO337 coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO337.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO337 in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO337 polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO337 DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO337 may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (*e.g.* Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO337 can be disrupted by various physical or chemical means, such as freeze-

thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO337 from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO337. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO337 produced.

E. Uses for PRO337

Nucleotide sequences (or their complement) encoding PRO337 have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO337 nucleic acid will also be useful for the preparation of PRO337 polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO337 (SEQ ID NO: 2 gene, or portions thereof (*e.g.*, without the N-terminal signal peptides, residues 1 to about 28), may be used as hybridization probes for a cDNA library to isolate the full-length gene or to isolate still other genes (for instance, those encoding naturally-occurring variants of PRO337 or PRO337 from other species) which have a desired sequence identity to the PRO337 sequence disclosed in Fig. 2 (SEQ ID NO: 2). Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from the nucleotide sequence of SEQ ID NO: 2 or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO337. By way of example, a screening method will comprise isolating the coding region of the PRO337 gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO337 gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

The ESTs disclosed in the present application (e.g., DNA28761, DNA42301) may similarly be employed as probes, using the methods disclosed herein. Alternatively, then can be used as a template from which to make probes.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO337 sequences.

Nucleotide sequences encoding a PRO337 can also be used to construct hybridization probes for mapping the gene which encodes that PRO337 and for the genetic analysis of individuals with genetic disorders. The

nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO337 encode a protein which binds to another protein (example, where the PRO337 is a receptor), the PRO337 can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO337 can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO337 or a receptor for PRO337. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO337 or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO337 can be used to clone genomic DNA encoding PRO337 in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO337. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO337 transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO337 introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO337. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Non-human homologues of PRO337 can be used to construct a PRO337 "knock out" animal which has a defective or altered gene encoding PRO337 as a result of homologous recombination between the endogenous gene encoding PRO337 and altered genomic DNA encoding PRO337 introduced into an embryonic cell of the animal. For example, cDNA encoding PRO337 can be used to clone genomic DNA encoding PRO337 in accordance with established techniques. A portion of the genomic DNA encoding PRO337 can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several

kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., Cell, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO337 polypeptide.

Neurotrimin as well as other members of the IgLON subfamily of the immunoglobulin superfamily have been identified to have effect upon neural patterning, differentiation, maturation and growth. As a result, PRO337 the human neurotrimin homolog polypeptides would be expected to have utility in diseases which are characterized by neural disfunction. For example, motoneuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, and various conditions involving spinal muscular atrophy, or paralysis. NGF variant formulations of the invention can be used to treat human neurodegenerative disorders, such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's Syndrome, nerve deafness, and Meniere's disease. Moreover PRO337 polypeptide may also be used as a cognitive enhancer, to enhance learning particularly in dementia or trauma, such as those associated with the above diseases

Further, PRO337 may be employed to treat neuropathy, and especially peripheral neuropathy. "Peripheral neuropathy" refers to a disorder affecting the peripheral nervous system, most often manifested as one or a combination of motor, sensory, sensorimotor, or autonomic neural dysfunction. The wide variety of morphologies exhibited by peripheral neuropathies can each be attributed uniquely to an equally wide number of causes. For example, peripheral neuropathies can be genetically acquired, can result from a systemic disease, or can be induced by a toxic agent. Examples include but are not limited to diabetic peripheral neuropathy, distal sensorimotor neuropathy, or autonomic neuropathies such as reduced motility of the gastrointestinal tract or atony of the urinary bladder. Examples of neuropathies associated with systemic disease include post-polio syndrome or AIDS-associated neuropathy; examples of hereditary neuropathies include Charcot-Marie-Tooth disease, Refsum's disease, Abetalipoproteinemia, Tangier disease, Krabbe's disease, Metachromatic leukodystrophy, Fabry's disease, and Dejerine-Sottas syndrome; and examples of neuropathies caused by a toxic agent include those caused by treatment with a chemotherapeutic agent such as vincristine, cisplatin, methotrexate, or 3'-azido-3'-deoxythymidine.

Correspondingly, neurotrimin antagonists would be expected to have utility in diseases characterized by excessive neuronal activity.

F. Anti-PRO337 Antibodies

The present invention further provides anti-PRO337 antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

The anti-PRO337 antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO337 polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

2. Monoclonal Antibodies

The anti-PRO337 antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO337 polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Rockville, Maryland. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the

production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO337. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., supra] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Humanized Antibodies

The anti-PRO337 antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)].

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO337, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production

of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

G. Uses for anti-PRO337 Antibodies

The anti-PRO337 antibodies of the invention have various utilities. For example, anti-PRO337 antibodies may be used in diagnostic assays for PRO337, *e.g.*, detecting its expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^{3}H, ^{14}C, ^{32}P, ^{35}S, or ^{125}I, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO337 antibodies also are useful for the affinity purification of PRO337 from recombinant cell culture or natural sources. In this process, the antibodies against PRO337 are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO337 to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO337, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO337 from the antibody.

Anti-PRO337 antibodies may also have utility in the treatment of diseases similar to those identified for PRO337 polypeptides. Anti-PRO337 antibodies would be especially useful for the treatment of diseases characterized by excessive neuronal activity and growth.

****

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way. All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Rockville, Maryland.

EXAMPLE 1

Isolation of cDNA clones Encoding Human PRO337

1. Preparation of oligo dT primed cDNA library

mRNA was isolated from human fetal brain tissue using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linkered cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

2. Preparation of random primed cDNA library

A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was

sized to 500-1000 bp, linkered with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37EC for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37EC). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, e.g. CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL^+, SUC^+, GAL^+.

Transformation was performed based on the protocol outlined by Gietz et al., Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30EC. The YEPD broth was prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about 2 x 10^6 cells/ml (approx. OD_{600}=0.1) into fresh YEPD broth (500 ml) and regrown to 1 x 10^7 cells/ml (approx. OD_{600}=0.4-0.5).

The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li_2OOCCH_3), and resuspended into LiAc/TE (2.5 ml).

Transformation took place by mixing the prepared cells (100 µl) with freshly denatured single stranded salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1 µg, vol. < 10 µl) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 µl, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li_2OOCCH_3, pH 7.5) was added. This mixture was gently mixed and incubated at 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500 µl, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells were then diluted into TE (1 ml)

and aliquots (200 µl) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Transformants were grown at 30°C for 2-3 days.

The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely et al., Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

4. Isolation of DNA by PCR Amplification

When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30 µl) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5 µl) was used as a template for the PCR reaction in a 25 µl volume containing: 0.5 µl Klentaq (Clontech, Palo Alto, CA); 4.0 µl 10 mM dNTP's (Perkin Elmer-Cetus); 2.5 µl Kentaq buffer (Clontech); 0.25 µl forward oligo 1; 0.25 µl reverse oligo 2; 12.5 µl distilled water. Appropriate forward and reverse oligonucleotides are used so as to anneal to the ADH promoter region and the amylase region of the vector, respectively, and which amplify a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

PCR is performed as follows:

a.	Denature		92°C,	5 minutes
b.	3 cycles of Denature		92°C,	30 seconds
		Anneal	59°C,	30 seconds
		Extend	72°C,	60 seconds
c.	3 cycles of Denature		92°C,	30 seconds
		Anneal	57°C,	30 seconds
		Extend	72°C,	60 seconds

d.	25 cycles of	Denature	92°C,	30 seconds
		Anneal	55°C,	30 seconds
		Extend	72°C,	60 seconds
e.		Hold	4°C	

Following the PCR, an aliquot of the reaction (5 µl) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook et al., supra. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA). An EST sequence isolated using the techniques described in this example is identified in Figure 9 (DNA 42301).

EXAMPLE 2

Isolation of cDNA clones Encoding Human PRO337

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996); http://blast.wustl/edu/blast/README.html] as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

A consensus DNA sequence encoding at least a portion PRO337 was assembled using phrap. In some cases, the consensus DNA sequence was extended using repeated cycles of blast and phrap to extend the consensus sequence as far as possible using the three sources of EST sequences listed above.

Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence. Forward and reverse PCR primers (notated as *.f and *.r, respectively) may range from 20 to 30 nucleotides (typically about 24), and are designed to give a PCR product of 100-1000 bp in length. The probe sequences (notated as *.p) are typically 40-55 bp (typically about 50) in length. These probes are identified in Figure 4 as OLI620 (28761.p), OLI621 (28761.f) and OLI622 (28761.r). In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification, as per *Ausubel et al., Current Protocols in Molecular Biology*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest by the in vivo cloning procedure suing the probe oligonucleotide and one of the

PCR primers.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO337 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA; Clontech, etc.) The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of DNA43316 is shown in Figure 1 (SEQ ID NO: 1). Clone DNA43316 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 134-136 [Kozak et al., supra] (Fig. 1; SEQ ID NO: 1). The predicted polypeptide precursor is 344 amino acids long. Clone DNA43316 has been deposited with ATCC and is assigned ATCC deposit no. 209487

Based on a BLAST-2 and FastA sequence alignment analysis of the full-length sequence, PRO337 shows amino acid sequence identity to rat neurotrimin (97%).

EXAMPLE 3

Use of PRO337 as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO337 as a hybridization probe.

DNA comprising the coding sequence of PRO337 (as shown in Figure 2, SEQ ID NO: 2) is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO337) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO337-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO337 can then be identified using standard techniques known in the art.

EXAMPLE 4

Expression of PRO337 in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO337 by recombinant expression in *E. coli*.

The DNA sequence encoding PRO337 (SEQ ID NO: 2) is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO337 coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO337 protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

EXAMPLE 5

Expression of PRO337 in mammalian cells

This example illustrates preparation of a glycosylated form of PRO337 by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO337 DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO337 DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-PRO337.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 µg pRK5-PRO337 DNA is mixed with about 1 µg DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 µl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M $CaCl_2$. To this mixture is added, dropwise, 500 µl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM $NaPO_4$, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μCi/ml ^{35}S-cysteine and 200 μCi/ml ^{35}S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO337 polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO337 may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μg pRK5-PRO337 DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μg/ml bovine insulin and 0.1 μg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO337 can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO337 can be expressed in CHO cells. The pRK5-PRO337 can be transfected into CHO cells using known reagents such as $CaPO_4$ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S-methionine. After determining the presence of PRO337 polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO337 can then be concentrated and purified by any selected method.

Epitope-tagged PRO337 may also be expressed in host CHO cells. The PRO337 may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO337 insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO337 can then be concentrated and purified by any selected method, such as by Ni^{2+}-chelate affinity chromatography.

EXAMPLE 6

Expression of PRO337 in Yeast

The following method describes recombinant expression of PRO337 in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO337 from the ADH2/GAPDH promoter. DNA encoding PRO337, a selected signal peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO337. For

secretion, DNA encoding PRO337 can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO337.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO337 can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO337 may further be purified using selected column chromatography resins.

EXAMPLE 7

Expression of PRO337 in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO337 in Baculovirus-infected insect cells.

The PRO337 is fused upstream of an epitope tag contained with a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO337 or the desired portion of the PRO337 (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression is performed as described by O'Reilley et al., Baculovirus expression vectors: A laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO337 can then be purified, for example, by Ni^{2+}-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM $MgCl_2$; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 Fm filter. A Ni^{2+}-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH

6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni^{2+}-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His_{10}-tagged PRO337 are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO337 can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

EXAMPLE 8

Preparation of Antibodies that Bind PRO337

This example illustrates preparation of monoclonal antibodies which can specifically bind PRO337.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO337, fusion proteins containing PRO337, and cells expressing recombinant PRO337 on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO337 immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect PRO337 antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO337. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO337. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO337 is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO337 monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

* * * *

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

Material	ATCC Dep. No.	Deposit Date
DNA43316-1237	209487	11/21/97

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposit will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC '122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

What is claimed is:

1. Isolated nucleic acid comprising DNA having at least a 95% sequence identity to (a) a DNA molecule encoding a PRO337 polypeptide comprising the sequence of amino acids about 29 to 344 of Fig. 2 (SEQ ID NO: 2).

2. The isolated nucleic acid of claim 1 comprising DNA having at least a 95% sequence identity to (a) a DNA molecule encoding a PRO337 polypeptide comprising the sequence of amino acids 1 to 344 of Fig. 2 (SEQ ID NO: 2), or (b) the complement of the DNA molecule of (a).

3. An isolated nucleic acid comprising DNA having at least a 95% sequence identity to (a) a DNA molecule encoding the same mature polypeptide encoded by the cDNA in ATCC Deposit No. 209487 (designation: DNA43316-1237), or (b) the complement of the DNA molecule of (a).

4. The isolated nucleic acid of claim 3 comprising the PRO337 encoding sequence of the cDNA in ATCC deposit No. (designation: DNA43316-1237), or a sequence which hybridizes thereto under stringent conditions.

5. A vector comprising the nucleic acid of claim 1.

6. The vector of claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

7. A host cell comprising the vector of claim 6.

8. The host cell of claim 7 wherein said cell is mammalian.

9. The host cell of claim 8 wherein said cell is a CHO cell.

10. The host cell of claim 7 wherein said cell is procaryotic.

11. The host cell of claim 10 wherein said cell is an *E. coli*.

12. The host cell of claim 7 wherein said cell is a yeast cell.

13. The host cell of claim 12 wherein said cell is *Saccharomyces cerevisiae*.

14. A process for producing PRO337 polypeptides comprising culturing the host cell of claim 7 under conditions suitable for expression of PRO337 and recovering PRO337 from the cell culture.

15. Isolated native sequence PRO337 polypeptide comprising amino acid residues about 29 to 344 of Fig. 2 (SEQ ID NO: 2).

16. Isolated native sequence PRO337 polypeptide comprising amino acid residues 1 to 344 of Fig. 2. (SEQ ID NO: 2).

17. Isolated native sequence PRO337 polypeptide encoded by the nucleotide deposited under accession number ATCC 209487.

18. A chimeric molecule comprising PRO337 polypeptide fused to a heterologous amino acid sequence.

19. The chimeric molecule of claim 17 wherein said heterologous amino acid sequence is an epitope tag sequence.

20. The chimeric molecule of claim 17 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

21. An antibody which specifically binds to human neurotrimin polypeptide.

22. The antibody of claim 21 wherein said antibody is a monoclonal antibody.

23. An isolated nucleic acid comprising the nucleotide sequence of DNA28761 (SEQ ID NO:3), DNA42301 (SEQ ID NO: 7) or their complements.

(SEQ ID NO: 1)

```
GTTGTGTCCT TCAGCAAAAC AGTGGATTTA AATCTCCTTG CACAAGCTTG AGAGCAACAC AATCTATCAG GAAAGAAAGA AAGAAAAAAA CCGAACCTGA 100
CAAAAAAGAA GAAAAAGAAG AAGAAAAAAA ATCATGAAAA CCATCCAGCC AAAAATGCAC AATTCTATCT CTTGGGCAAT CTTCACGGGG CTGGCTGCTC 200
TGTGTCTCTT CCAAGGAGTG CCCGTGCGCA GCGGAGATGC CACCTTCCCC AAAGCTATGG ACAACGTGAC GGTCCGGCAG GGGGAGAGCG CCACCCTCAG 300
GTGCACTATT GACAACCGGG TCACCCGGGT GGCCTGGCTA AACCGCAGCA CCATCCTCTA TGCTGGGAAT GACAAGTGGT GCCTGGATCC TCGCGTGGTC 400
CTTCTGAGCA ACACCCAAAC GCAGTACAGC ATCGAGATCC AGAACGTGGA TGTGTATGAC GAGGGCCCTT ACACCTGCTC GGTGCAGACA GACAACCACC 500
CAAAGACCTC TAGGGTCCAC CTCATTGTGC AAGTATCTCC CAAAATTGTA GAGATTTCTT CAGATATCTC CATTAATGAA GGGAACAATA TTAGCCTCAC 600
CTGCATAGCA ACTGGTAGAC CAGAGCCTAC GGTTACTTGG AGACACATCT CTCCCAAAGC GGTTGGCTTT GTGAGTGAAG ACGAATACTT GGAAATTCAG 700
GGCATCACCC GGGAGCAGTC AGGGGACTAC GAGTGCAGTG CCTCCAATGA CGTGGCCGCG CCCGTGGTAC GGAGAGTAAA GGTCACCGTG AACTATCCAC 800
CATACATTTC AGAAGCCAAG GGTACAGGTG TCCCCGTGGG ACAAAAGGGG ACACTGCAGT GTGAAGCCTC AGCAGTCCCC TCAGCAGAAT TCCAGTGGTA 900
CAAGGATGAC AAAAGACTGA TTGAAGGAAA GAAAGGGGTG AAAGTGGAAA ACAGACCTTT CCTCTCAAAA CTCATCTTCT TCAATGTCTC TGAACATGAC 1000
TATGGGAACT ACACTTGCGT GGCCTCCAAC AAGCTGGGCC ACACCAATGC CAGCATCATG CTATTTGGTC CAGGCGCCGT CAGCGAGGTG AGCAACGGCA 1100
CGTCGAGGAG GGCAGGCTGC GTCTGGCTGC TGCCTCTTCT GGTCTTGCAC CTGCTTCTCA AATTTTGATG TGAGTGCCAC TTCCCCACCC GGGAAAGGCT 1200
GCCGCCACCA CCACCACCAA CACAACAGCA ATGGCAACAC CGACAGCAAC CAATCAGATA TATACAAATG AAATTAGAAG AAACACAGCC TCATGGGACA 1300
GAAATTTGAG GGAGGGGAAC AAAGAATACT TTGGGGGGAA AAGAGTTTTA AAAAAGAAAT TGAAAATTGC CTTGCAGATA TTTAGGTACA ATGGAGTTTT 1400
CTTTTCCCAA ACGGGAAGAA CACAGCACAC CCGGCTTGGA CCCACTGCAA GCTGCATCGT GCAACCTCTT TGGTGCCAGT GTGGGCAAGG GCTCAGCCTC 1500
TCTGCCCACA GAGTGCCCCC ACGTGGAACA TTCTGGAGCT GGCCATCCCA AATTCAATCA GTCCATAGAG ACGAACAGAA TGAGACCTTC CGGCCCAAGC 1600
GTGGCGCTGC GGGCACTTTG GTAGACTGTG CCACCACGGC GTGTGTTGTG AAACGTGAAA TAAAAAGAGC AAAAAAAAA 1679
```

Figure 1

(SEQ ID NO. 2)

Met Lys Thr Ile Gln Pro Lys Met His Asn Ser Ile Ser Trp Ala
1 5 10 15

Ile Phe Thr Gly Leu Ala Ala Leu Cys Leu Phe Gln Gly Val Pro
20 25 30

Val Arg Ser Gly Asp Ala Thr Phe Pro Lys Ala Met Asp Asn Val
35 40 45

Thr Val Arg Gln Gly Glu Ser Ala Thr Leu Arg Cys Thr Ile Asp
50 55 60

Asn Arg Val Thr Arg Val Ala Trp Leu Asn Arg Ser Thr Ile Leu
65 70 75

Tyr Ala Gly Asn Asp Lys Trp Cys Leu Asp Pro Arg Val Val Leu
80 85 90

Leu Ser Asn Thr Gln Thr Gln Tyr Ser Ile Glu Ile Gln Asn Val
95 100 105

Asp Val Tyr Asp Glu Gly Pro Tyr Thr Cys Ser Val Gln Thr Asp
110 115 120

Asn His Pro Lys Thr Ser Arg Val His Leu Ile Val Gln Val Ser
125 130 135

Pro Lys Ile Val Glu Ile Ser Ser Asp Ile Ser Ile Asn Glu Gly
140 145 150

Asn Asn Ile Ser Leu Thr Cys Ile Ala Thr Gly Arg Pro Glu Pro
155 160 165

Thr Val Thr Trp Arg His Ile Ser Pro Lys Ala Val Gly Phe Val
170 175 180

Ser Glu Asp Glu Tyr Leu Glu Ile Gln Gly Ile Thr Arg Glu Gln
185 190 195

Ser Gly Asp Tyr Glu Cys Ser Ala Ser Asn Asp Val Ala Ala Pro
200 205 210

Val Val Arg Arg Val Lys Val Thr Val Asn Tyr Pro Pro Tyr Ile
215 220 225

Ser Glu Ala Lys Gly Thr Gly Val Pro Val Gly Gln Lys Gly Thr
230 235 240

Leu Gln Cys Glu Ala Ser Ala Val Pro Ser Ala Glu Phe Gln Trp
245 250 255

Tyr Lys Asp Asp Lys Arg Leu Ile Glu Gly Lys Lys Gly Val Lys
260 265 270

Val Glu Asn Arg Pro Phe Leu Ser Lys Leu Ile Phe Phe Asn Val
275 280 285

Ser Glu His Asp Tyr Gly Asn Tyr Thr Cys Val Ala Ser Asn Lys
290 295 300

Leu Gly His Thr Asn Ala Ser Ile Met Leu Phe Gly Pro Gly Ala
305 310 315

Val Ser Glu Val Ser Asn Gly Thr Ser Arg Arg Ala Gly Cys Val
320 325 330

Trp Leu Leu Pro Leu Leu Val Leu His Leu Leu Leu Lys Phe
335 340 344

Figure 2

```
HSC0PH101      1 CAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTA
(SEQ ID NO: 4)
   R18841      1  AAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTA
(SEQ ID NO: 5)
<DNA28761>     1 CAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTA
(SEQ ID NO: 3)

HSC0PH101     51 GAGATTTCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCAC
   R18841     50 GAGATTTCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCAC
<DNA28761>    51 GAGATTTCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCAC

HSC0PH101    101 CTGCATAGCAACTGGTAGACCAGAGCCTACGGTTACTTGGAGACACATCT
   R18841    100 CTGCATAGCAACTGGTAGACCAGAGCCTACGGTTACTTGGAGACACATCT
<DNA28761>   101 CTGCATAGCAACTGGTAGACCAGAGCCTACGGTTACTTGGAGACACATCT

HSC0PH101    151 CTCCCAAAGCGGTTGGCTTTGTGAGTGAAGACGAATACTTGGAAATTCAG
   R18841    150 CTCCCAAAGCGGTTGGCTTTGTGAGTGAAGACGAATACTTGGAAATTCAG
  1333422      1                                          GAAATTCAG
(SEQ ID NO: 6)
<DNA28761>   151 CTCCCAAAGCGGTTGGCTTTGTGAGTGAAGACGAATACTTGGAAATTCAG

HSC0PH101    201 GGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTCCAATGA
   R18841    200 GGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTCCAATGA
  1333422     10 GGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTCCAATGA
<DNA28761>   201 GGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTCCAATGA

HSC0PH101    251 CGTGG-CCGCGCCCGTGG-TACGGAG-AGTAAAGGTCACCGTG
   R18841    250 CGTGGGCCGCGCCCGTGGGTACGGAGGAGTAAAGGTCACCGTGAACTATT
  1333422     60 CGTGG-CCGCGCCCGTGG-TACGGAG-AGTAAAGGTCACCGTGAACTAT-
<DNA28761>   251 CGTGGGCCGCGCCCGTGGGTACGGAG-AGTAAAGGTCACCGTGAACTATT

   R18841    300 CCACCATACATTTCAGGAAGCCAAGGGTACAGGTGTCCCCGTGGGCAAAA
  1333422    110 CCACCATACATTTCAG-AAGCCAAGGGTACAGGTGTCCCCGTGGG-ACAA
<DNA28761>   301 CCACCATACATTTCAGGAAGCCAAGGGTACAGGTGTCCCCGTGGGCAAAA

   R18841    350 AAGGGGACACTTGCAGTGTTGAAGCCTTCAGCAGTTCCCTTNAGCAGGTT
  1333422    160 AAGGGGACACT-GCAGTGT-GAAGCCT-CAGCAGTCCCCT-CAGCAGAAT
<DNA28761>   351 AAGGGGACACTTGCAGTGTTGAAGCCTTCAGCAGTTCCCTTNAGCAGGTT

   R18841    400 TTCCGTTGGTTACAAG
  1333422    210 TCCAGT-GGT-ACAAGGA
<DNA28761>   401 TTCCGTTGGTTACAAG
```

List of clones used in Assembly:

EST:	HSC0PH101
	R18841
Incyte:	1333422

Figure 3

OLI620 28761.p
(SEQ ID NO: 8)

AGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGACGAATACTT

Figure 4A

OLI621 28761.f
(SEQ ID NO: 9)

GGGAACAATATTAGCCTCACCTGC

Figure 4B

OLI622 28761.r
(SEQ ID NO: 10)

GGACACCTGTACCCTTGGCTTC

Figure 4C

		Score	Match	Pct
I56551	neurotrimin - rat	1637	309	97
CEPU_CHICK	Cepu-1 protein precursor - gallus gallus	1417	261	85
G55A_CHICK	Neurite inhibitor gp55a - gallus gallus, C-te...	1058	201	73
LAMP_HUMAN	Limbic system-associated membrane protein pre...	750	132	59
OPCM_HUMAN	Opioid binding protein/cell adhesion molecule...	682	125	84

I56551 neurotrimin - rat (344 aa)
Score = 1637 (750.2 bits), Expect = 1.2e-221, P = 1.2e-221
Identities = 309/317 (97%), Positives = 315/317 (99%) at 28,28

```
DNA43316   28 GVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNRVTRVAWLNRSTILYAGNDKWCLDPR
              ************************************************************
  I56551   28 GVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNRVTRVAWLNRSTILYAGNDKWCLDPR
```

(SEQ ID NO: 11)

```
DNA43316   88 VVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPKTSRVHLIVQVSPKIVEISSDISI
              ************************************************************
  I56551   88 VVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPKTSRVHLIVQVSPKIVEISSDISI

DNA43316  148 NEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYLEIQGITREQSGDYECSASNDV
              ***********************************************.**********
  I56551  148 NEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYLEIQGITREQSGEYECSASNDV

DNA43316  208 AAPVVRRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAVPSAEFQWYKDDKRLIEGKK
              ******** ***************************************.******.****
  I56551  208 AAPVVRRVNVTVNYPPYISEAKGTGVPVGQKGTLQCEASAVPSAEFQWFKDDKRLVEGKK

DNA43316  268 GVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIMLPGPGAVSEVSNGTSRRA
              ***********.* ***************************************.*******
  I56551  268 GVKVENRPFLSRLTFFNVSEHDYGNYTCVASNKLGHTNASIMLPGPGAVSEVNNGTSRRA

DNA43316  328 GCVWLLPLLVLHLLLKF
              **.***************
  I56551  328 GCIWLLPLLVLHLLLKF
```

Figure 5A

CEPU_CHICK Cepu-1 protein precursor - gallus gallus (353 aa)
Score = 1417 (649.3 bits), Expect = 1.3e-200, Sum P(2) = 1.3e-200
Identities = 261/307 (85%), Positives = 288/307 (93%) at 5,3

```
DNA43316      5 QPKMHNSISWAIFTGLAALCLFQGVPVRSGDATPPKAMDNVTVRQGESATLRCTIDNRVT
                * ** . .** ** *.*** ********************************..*****
CEPU_CHICK    3 QAKMQHPVSWVIFAGMAALLLFQGVPVRSGDATPPKAMDNVTVRQGESATLRCSVDNRVT
(SEQ ID NO 12)

DNA43316     65 RVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPK
                ********.******************.**.*****.* .*********************
CEPU_CHICK   63 RVAWLNRSSILYAGNDKWCLDPRVVLLANTKTQYSIQIHDVDVYDEGPYTCSVQTDNHPK

DNA43316    125 TSRVHLIVQVSPKIVEISSDISINEGGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDE
                *************** * ********* *.**********.**.************.****
CEPU_CHICK  123 TSRVHLIVQVSPKITETSSDISINEGGGNVSLTCIATGRPDPTITWRHISPKAVGFISEDE

DNA43316    185 YLEIQGITREQSGDYECSASNDVAAPVVRRVKVTVNYPPYISEAKGTGVPVGQKGTLQCE
                **** ********.**************.*************.** ********* * **
CEPU_CHICK  183 YLEITGITREQSGEYECSASNDVAAPVVQRVKVTVNYPPYISDAKSTGVPVGQKGILMCE

DNA43316    245 ASAVPSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHT
                ********.********** **.**.****. * *.* ****** ***********.**.*
CEPU_CHICK  243 ASAVPSADFQWYKDDKRLAEGQKGLKVENKAFFSRLTFFNVSEQDYGNYTCVASNQLGNT

DNA43316    305 NASIMLF
                ***..*.
CEPU_CHICK  303 NASMILY
```

Score = 76 (34.8 bits), Expect = 1.3e-200, Sum P(2) = 1.3e-200
Identities = 15/33 (45%), Positives = 20/33 (60%) at 312,321

```
DNA43316    312 GPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF
                ***** . ..*  **  * *** * .  *  .*
CEPU_CHICK  321 GPGAVHDGNSGAWRRGSCAWLLALPLAQLARQF
```

Figure 5B

G55A_CHICK Neurite inhibitor gp55a - gallus gallus, C-terminal fra (274 aa)
Score = 1058 (484.8 bits), Expect = 1.1e-141, P = 1.1e-141
Identities = 201/274 (73%), Positives = 230/274 (83%) at 71,1

```
DNA43316     71 RSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPKTSRVHL
                ************* .* ***.****.*****.* ****************************
G55A_CHICK    1 RSTILYAGNDKWSIDNRVVILSNTKTQYSIKIHNVDVYDEGPYTCSVQTDNHPKTSRVHL
(SEQ ID NO: 13)

DNA43316    131 IVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYLEIQG
                **** *.** *****..***....* *.* ***********.* *  ********** *
G55A_CHICK   61 IVQVPPQIVNISSDITVNEGSSVTLMCLAFGRPEPTVTWRHLSGKGQGFVSEDEYLEITG

DNA43316    191 ITREQSGDYECSASNDVAAPVVRRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAVPS
                *******.***** **** * **.************ ** **  ***** *********
G55A_CHICK  121 ITREQSGEYECSAVNDVAVPDVRKVKVTVNYPPYISNAKNTGASVGQKGILQCEASAVPV

DNA43316    251 AEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIML
                *****.*.* **  * .**..*..  ** * ****** *********.****.*****.*
G55A_CHICK  181 AEFQWFKEDTRLANGLEGVRIESKGRLSTLTFFNVSEKDYGNYTCVATNKLGNTNASIIL

DNA43316    311 FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF
                .***** .  *  ** *  .  *   *. *** *
G55A_CHICK  241 YGPGAVHDSGNAASRAAAGLCLWATLLARLLLDF
```

LAMP_HUMAN Limbic system-associated membrane protein precursor - (338 aa)
Score = 750 (343.7 bits), Expect = 1.7e-123, Sum P(3) = 1.7e-123
Identities = 132/225 (58%), Positives = 178/225 (79%) at 38,34

```
DNA43316     38 FPKAMDNVTVRQGESATLRCTIDNRVTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQ
                *  . **.*****..* *** .... ..******* *..**.*** ***** *      .
LAMP_HUMAN   34 FNRGTDNITVRQGDTAILRCVLEDKNSKVAWLNRSGIIFAGHDKWSLDPRVELEKRHSLE
(SEQ ID NO 14)

DNA43316     98 YSIEIQNVDVYDEGPYTCSVQTDNHPKTSRVHLIVQVSPKIVEISSDISINEGNNISLTC
                **. ** ******* ******* . ****.*.***** ***  ****...***.*..* *
LAMP_HUMAN   94 YSLRIQKVDVYDEGSYTCSVQTQHEPKTSQVYLIVQVPPKISNISSDVTVNEGSNVTLVC

DNA43316    158 IATGRPEPTVTWRHISPKAVGFVSEDEYLEIQGITREQSGDYECSASNDVAAPVVRRVKV
                .* ***** .****..*    *  *.***** ******** *** *.*.*..  *..***
LAMP_HUMAN  154 MANGRPEPVITWRHLTPTGREFEGEEEYLEILGITREQSGKYECKAANEVSSADVKQVKV

DNA43316    218 TVNYPPYISEAKGTGVPVGQKGTLQCEASAVPSAEFQWYKDDKRL
                ****** *.*.*      *.. .*.*******. .*.**.** *.
LAMP_HUMAN  214 TVNYPPTITESKSNEATTGRQASLKCEASAVPAPDFEWYRDDTRI
```

Figure 5C

Score = 149 (68.3 bits), Expect = 1.7e-123, Sum P(3) = 1.7e-123
Identities = 30/66 (45%), Positives = 42/66 (63%) at 259,254

```
DNA43316    259 DKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIMLFGPGAVSE
                *  *   *.....   * *   **.*  ********.**** ****..** **.*
LAMP_HUMAN  254 DDTRINSANGLEIKSTEGQSSLTVTNVTEEHYGNYTCVAANKLGVTNASLVLPRPGSVRG

DNA43316    319 VSNGTS
                ..  *
LAMP_HUMAN  314 INGSIS
```

Score = 35 (16.0 bits), Expect = 1.7e-123, Sum P(3) = 1.7e-123
Identities = 10/23 (43%), Positives = 13/23 (56%) at 321,315

```
DNA43316    321 NGTSRRAGCVWLLPLLVLHLLLK
                **.   *  .***   .* ** *
LAMP_HUMAN  315 NGSISLAVPLWLLAASLLCLLSK
```

OPCM_HUMAN Opioid binding protein/cell adhesion molecule precurso (345 aa)
Score = 682 (312.5 bits), Expect = 5.3e-171, Sum P(2) = 5.3e-171
Identities = 125/148 (84%), Positives = 139/148 (93%) at 28,28

```
DNA43316     28 GVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNRVTRVAWLNRSTILYAGNDKWCLDPR
                ***********************************.********************** .***
OPCM_HUMAN   28 GVPVRSGDATFPKAMDNVTVRQGESATLRCTIDDRVTRVAWLNRSTILYAGNDKWSIDPR
(SEQ ID NO 15)

DNA43316     88 VVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPKTSRVHLIVQVSPKIVEISSDISI
                *..* ** ***** ****************************** *.*. *****..
OPCM_HUMAN   88 VIILVNTPTQYSIMIQNVDVYDEGPYTCSVQTDNHPKTSRVHLIVQVPPQIMNISSDITV

DNA43316    148 NEGNNISLTCIATGRPEPTVTWRHISPK
                ***....* *.* ***********.* *
OPCM_HUMAN  148 NEGSSVTLLCLAIGRPEPTVTWRHLSVK
```

Score = 596 (273.1 bits), Expect = 5.3e-171, Sum P(2) = 5.3e-171
Identities = 112/167 (67%), Positives = 132/167 (79%) at 178,179

```
DNA43316    178 GFVSEDEYLEIQGITREQSGDYECSASNDVAAPVVRRVKVTVNYPPYISEAKGTGVPVGQ
                ***********  * *.***.***** ****** **.**.*********.** *** ***
OPCM_HUMAN  179 GFVSEDEYLEISDIKRDQSGEYECSALNDVAAPDVRKVKITVNYPPYISKAKNTGVSVGQ

DNA43316    238 KGTLQCEASAVPSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVA
                ** * ******* *****.*..  **   *   *...**.   .* * ****** *********
OPCM_HUMAN  239 KGILSCEASAVPMAEFQWFKEETRLATGLDGMRIENKGRMSTLTFFNVSEKDYGNYTCVA

DNA43316    298 SNKLGHTNASIMLFGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF
                .****.***** *.*****  .   *  **   *.**   *. *  .**
OPCM_HUMAN  299 TNKLGNTNASITLYGPGAVIDGVNSASRALACLWLSGTLLAHFFIKF
```

Figure 5D

```
> /usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43316 (1795 bases)
> length: 1679 bp (circular)

  1 GTTGTGTCCT TCAGCAAAAC AGTGGATTTA AATCTCCTTG CACAAGCTTG AGAGCAACAC AATCTATCAG GAAAGAAAGA AAGAAAAAAA CCGAACCTGA
    CAACACAGGA AGTCGTTTTG TCACCTAAAT TTAGAGGAAC GTGTTCGAAC TCTCGTTGTG TTAGATAGTC CTTTCTTTCT TTCTTTTTTT GGCTTGGACT

101 CAAAAAAGAA GAAAAAGAAG AAGAAAAAAA ATCATGAAAA CCATCCAGCC AAAAATGCAC AATTCTATCT CTTGGGCAAT CTTCACGGGG CTGGCTGCTC
    GTTTTTTCTT CTTTTTCTTC TTCTTTTTTT TAGTACTTTT GGTAGGTCGG TTTTTACGTG TTAAGATAGA GAACCCGTTA GAAGTGCCCC GACCGACGAG
  1                                   M K T   I Q P   K N H   N S I S   W A I   P T G   L A A L
                                      ^Met

201 TGTGTCTCTT CCAAGGAGTG CCCGTGCGCA GCGGAGATGC CACCTTCCCC AAAGCTATGG ACAACGTGAC GGTCCGGCAG GGGGAGAGCG CCACCCTCAG
    ACACAGAGAA GGTTCCTCAC GGGCACGCGT CGCCTCTACG GTGGAAGGGG TTTCGATACC TGTTGCACTG CCAGGCCGTC CCCCTCTCGC GGTGGGAGTC
 24  C L P   Q G V   P V R S   G D A   T P P   K A M D   N V T   V R Q   G E S A   T L R

301 GTGCACTATT GACAACCGGG TCACCCGGGT GGCCTGGCTA AACCGCAGCA CCATCCTCTA TGCTGGGAAT GACAAGTGGT GCCTGGATCC TCGCGTGGTC
    CACGTGATAA CTGTTGGCCC AGTGGGCCCA CCGGACCGAT TTGGCGTCGT GGTAGGAGAT ACGACCCTTA CTGTTCACCA CGGACCTAGG AGCGCACCAG
 57 C T I   D N R V   T R V   A W L   N R S T   I L Y   A G N   D K W C   L D P   R V V

401 CTTCTGAGCA ACACCCAAAC GCAGTACAGC ATCGAGATCC AGAACGTGGA TGTGTATGAC GAGGGCCCTT ACACCTGCTC GGTGCAGACA GACAACCACC
    GAAGACTCGT TGTGGGTTTG CGTCATGTCG TAGCTCTAGG TCTTGCACCT ACACATACTG CTCCCGGGAA TGTGGACGAG CCACGTCTGT CTGTTGGTGG
 90 L L S N   T Q T   Q Y S   I E I Q   N V D   V Y D   E G P Y   T C S   V Q T   D N H P

501 CAAAGACCTC TAGGGTCCAC CTCATTGTGC AAGTATCTCC CAAAATTGTA GAGATTTCTT CAGATATCTC CATTAATGAA GGGAACAATA TTAGCCTCAC
    GTTTCTGGAG ATCCCAGGTG GAGTAACACG TTCATAGAGG GTTTTAACAT CTCTAAAGAA GTCTATAGAG GTAATTACTT CCCTTGTTAT AATCGGAGTG
124  K T S   R V H   L I V Q   V S P   K I V   E I S S   D I S   I N E   G N N I   S L T

601 CTGCATAGCA ACTGGTAGAC CAGAGCCTAC GGTTACTTGG AGACACATCT CTCCCAAAGC GGTTGGCTTT GTGAGTGAAG ACGAATACTT GGAAATTCAG
    GACGTATCGT TGACCATCTG GTCTCGGATG CCAATGAACC TCTGTGTAGA GAGGGTTTCG CCAACCGAAA CACTCACTTC TGCTTATGAA CCTTTAAGTC
157  C I A   T G R P   E P T   V T W   R H I S   P K A   V G F   V S E D   E Y L   E I Q

701 GGCATCACCC GGGAGCAGTC AGGGGACTAC GAGTGCAGTG CCTCCAATGA CGTGGCCGCG CCCGTGGTAC GGAGAGTAAA GGTCACCGTG AACTATCCAC
    CCGTAGTGGG CCCTCGTCAG TCCCCTGATG CTCACGTCAC GGAGGTTACT GCACCGGCGC GGGCACCATG CCTCTCATTT CCAGTGGCAC TTGATAGGTG
190 G I T R   E Q S   G D Y   E C S A   S N D   V A A   P V V R   R V K   V T V   N Y P P

801 CATACATTTC AGAAGCCAAG GGTACAGGTG TCCCCGTGGG ACAAAAGGGG ACACTGCAGT GTGAAGCCTC AGCAGTCCCC TCAGCAGAAT TCCAGTGGTA
    GTATGTAAAG TCTTCGGTTC CCATGTCCAC AGGGGCACCC TGTTTTCCCC TGTGACGTCA CACTTCGGAG TCGTCAGGGG AGTCGTCTTA AGGTCACCAT
224  Y I S   E A K   G T G V   P V G   Q K G   T L Q C   E A S   A V P   S A E F   Q W Y
```

Figure 6A

```
901 CAAGGATGAC AAAAGACTGA TTGAAGGAAA GAAAGGGGTG AAAGTGGAAA ACAGACCTTT CCTCTCAAAA CTCATCTTCT TCAATGTCTC TGAACATGAC
    GTTCCTACTG TTTTCTGACT AACTTCCTTT CTTTCCCCAC TTTCACCTTT TGTCTGGAAA GGAGAGTTTT GAGTAGAAGA AGTTACAGAG ACTTGTACTG
257  K  D  D   K  R  L  I   E  G  K   K  G  V   K  V  E  N   R  P  F  L  S  K   L  I  F  F   N  V  S   E  H  D

1001 TATGGGAACT ACACTTGCGT GGCCTCCAAC AAGCTGGGCC ACACCAATGC CAGCATCATG CTATTTGGTC CAGGCGCCGT CAGCGAGGTG AGCAACGGCA
     ATACCCTTGA TGTGAACGCA CCGGAGGTTG TTCGACCCGG TGTGGTTACG GTCGTAGTAC GATAAACCAG GTCCGCGGCA GTCGCTCCAC TCGTTGCCGT
290  Y  G  N  Y   T  C  V   A  S  N   K  L  G  H   T  N  A   S  I  M   L  F  G  P   G  A  V   S  E  V   S  N  G  T

1101 CGTCGAGGAG GGCAGGCTGC GTCTGGCTGC TGCCTCTTCT GGTCTTGCAC CTGCTTCTCA AATTCTGATG TGAGTGCCAC TTCCCCACCC GGGAAAGGCT
     GCAGCTCCTC CCGTCCGACG CAGACCGACG ACGGAGAAGA CCAGAACGTG GACGAAGAGT TTAAAACTAC ACTCACGGTG AAGGGGTGGG CCCTTTCCGA
324   S  R  R   A  G  C   V  W  L  L   P  L  L   V  L  H   L  L  L  K   F  O
                                                ^43316.tm.f1                ^43316.tm.p1

1201 GCCGCCACCA CCACCACCAA CACAACAGCA ATGGCAACAC CGACAGCAAC CAATCAGATA TATACAAATG AAATTAGAAG AAACACAGCC TCATGGGACA
     CGGCGGTGGT GGTGGTGGTT GTGTTGTCGT TACCGTTGTG GCTGTCGTTG GTTAGTCTAT ATATGTTTAC TTTAATCTTC TTTGTGTCGG AGTACCCTGT
              ^43316.tm.r1

1301 GAAATTTGAG GGAGGGGAAC AAAGAATACT TTGGGGGGAA AAGAGTTTTA AAAAAGAAAT TGAAAATTGC CTTGCAGATA TTTAGGTACA ATGGAGTTTT
     CTTTAAACTC CCTCCCCTTG TTTCTTATGA AACCCCCCTT TTCTCAAAAT TTTTTCTTTA ACTTTTAACG GAACGTCTAT AAATCCATGT TACCTCAAAA

1401 CTTTTCCCAA ACGGGAAGAA CACAGCACAC CCGGCTTGGA CCCACTGCAA GCTGCATCGT GCAACCTCTT TGGTGCCAGT GTGGGCAAGG GCTCAGCCTC
     GAAAAGGGTT TGCCCTTCTT GTGTCGTGTG GGCCGAACCT GGGTGACGTT CGACGTAGCA CGTTGGAGAA ACCACGGTCA CACCCGTTCC CGAGTCGGAG

1501 TCTGCCCACA GAGTGCCCCC ACGTGGAACA TTCTGGAGCT GGCCATCCCA AATTCAATCA GTCCATAGAG ACGAACAGAA TGAGACCTTC CGGCCCAAGC
     AGACGGGTGT CTCACGGGGG TGCACCTTGT AAGACCTCGA CCGGTAGGGT TTAAGTTAGT CAGGTATCTC TGCTTGTCTT ACTCTGGAAG GCCGGGTTCG

1601 GTGGCGCTGC GGGCACTTTG GTAGACTGTG CCACCACGGC GTGTGTTGTG AAACGTGAAA TAAAAAGAGC AAAAAAAAA
     CACCGCGACG CCCGTGAAAC CATCTGACAC GGTGGTGCCG CACACAACAC TTTGCACTTT ATTTTTCTCG TTTTTTTTT
```

Figure 6B

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA43316
<subunit 1 of 1, 344 aa, 1 stop
<MW: 37954, pI: 8.22, NX(S/T): 7
<     1       10        20        30        40        50        60        70
<     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
      MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNRVTRVAWLN

<     71      80        90        100       110       120       130       140
<     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
      RSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPKTSRVHLIVQVSPKIVE

<     141   150       160       170       180       190       200       210
<     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
      ISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYLEIQGITREQSGDYECSASNDVAAP

<     211   220       230       240       250       260       270       280
<     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
      VVRRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAVPSAEFQWYKDDKRLIEGKKGVKVENRPFLSKL

<     281   290       300       310       320       330       340
<     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
      IFFNVSEHDYGNYTCVASNKLGHTNASIMLFGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF
```

Figure 7A

Important features:

Signal peptide:
1-28 MKTIQPKMHNSISWAIFTGLAALCLFQG
Transmembrane domain:
none

Immunoglobulins and major histocompatibility domain:
113-\50 YTCSVQtDNHPKTsRVHL

Figure 7B

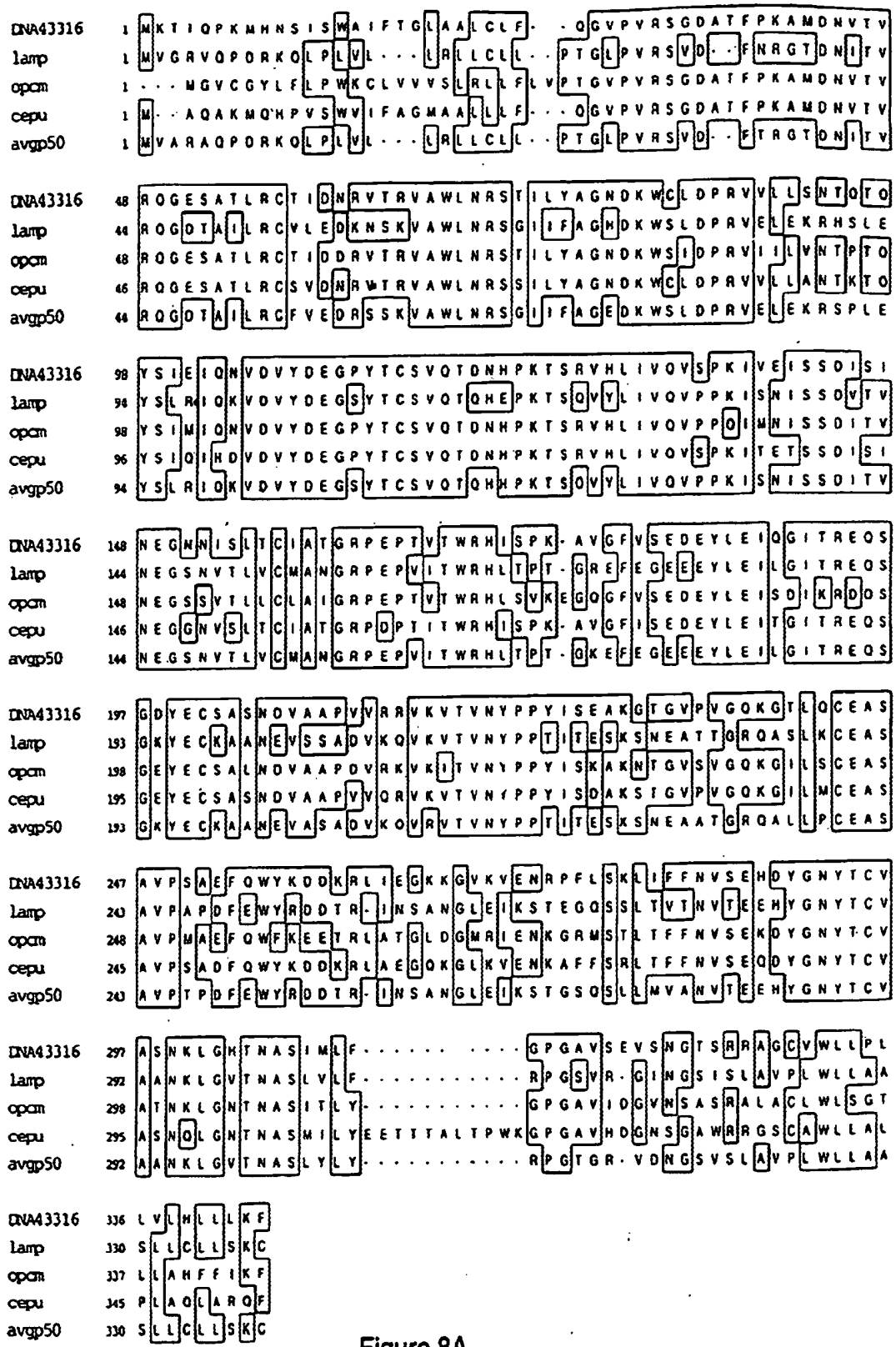

Figure 8A

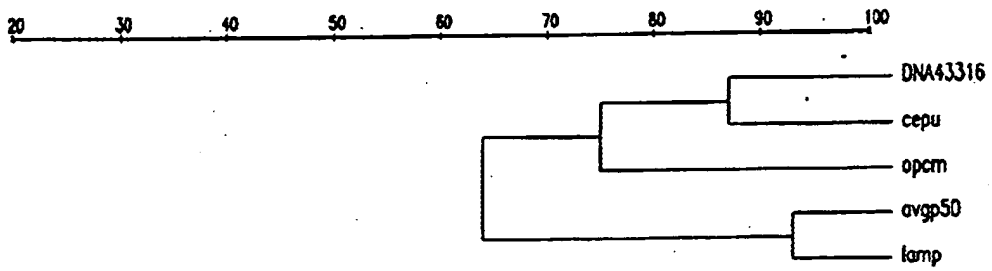

Figure 8B

```
DNA43316  123 GAAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGCAATCT
              ************************************************************
DNA42301    1 GAAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGCAATCT

DNA43316  183 TCACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCA
              ************************************************************
DNA42301   61 TCACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCA

DNA43316  243 CCTTCCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGT
              ************************************************************
DNA42301  121 CCTTCCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGT

DNA43316  303 GCACTATTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATG
              ************************************************************
DNA42301  181 GCACTATTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATG

DNA43316  363 CTGGGAATGACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGC
              ************************************************************
DNA42301  241 CTGGGAATGACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGC

DNA43316  423 AGTACAGCATCGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGG
              ************************************************************
DNA42301  301 AGTACAGCATCGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGG

DNA43316  483 TGCAGACAGACAACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCA
              ************************************************************
DNA42301  361 TGCAGACAGACAACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCA

DNA43316  543 AAATTGTAGAGATTTCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCT
              ************************************************************
DNA42301  421 AAATTGTAGAGATTTCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCT

DNA43316  603 GCATAGCAACTGGTAGACCAGAG
              ***********************
DNA42301  481 GCATAGCAACTGGTAGACCAGAG
```

Figure 9